Derivative Formula and Harnack Inequality for Jump Processes*

Feng-Yu Wang

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

and

Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

Email: wangfy@bnu.edu.cn; F.Y.Wang@swansea.ac.uk

January 20, 2013

Abstract

By using lower bound conditions of the Lévy measure, derivative formulae and Harnack inequalities are derived for linear stochastic differential equations driven by Lévy processes. As applications, explicit gradient estimates and heat kernel inequalities are presented. As byproduct, a new Girsanov theorem for Lévy processes is derived.

AMS subject Classification: 60J75, 60J45.
Keywords: Lévy process, derivative formula, gradient estimate, Harnack inequality.

1 Introduction

The derivative formula enables one to derive explicit gradient estimates; while the Harnack inequality has been applied to the study of heat kernel estimates, contractivity properties, transportation-const inequalities and properties of the invrainat probability measures, see e.g. [33, 9, 38] and references within (see §4.2 for some general results).

Recall that Bismut's derivative formula of elliptic diffusion semigroups [5], also known as Bismut-Elworthy-Li formula due to [10], is a powerful tool for stochastic analysis on Riemannian manifolds and has been extended and applied to SDEs (stochastic differential equations) driven by noises with a non-trivial Gaussian parts, see e.g. [30] and references

*Supported in part by WIMCS, SRFDP and the Fundamental Research Funds for the Central Universities.

within for the study of diffusion-jump processes. But, up to our knowledge, explicit derivative formula relying only on the Lévy measure is not yet available.

On the other hand, by using couplings constructed through Girsanov transforms, the dimension-free Harnack inequality, first introduced by the author in [31] for diffusion semigroups on manifolds, has been established and applied to various SDEs and SPDEs driven by Gaussian noises, see [2, 3, 9, 18, 20, 21, 25, 32, 33, 37, 36, 38, 40]. Since arguments used in these references essentially relies on special properties of the Brownian motion, they do not apply to the jump setting. Therefore, it is in particular interesting to built up a reasonable theory on derivative formula and Harnack inequality for pure jump processes.

In this paper, we aim to establish derivative formula and Harnack inequality for the semigroup associated to SDEs driven by Lévy jump processes using lower bound conditions of the Lévy measure. As observed in a recent paper [35], where the coupling property is confirmed for a class of linear SDEs driven by Lévy processes, the Mecke formula on the Poisson space will play an alternative role in the jump case to the Girsanov transform in the diffusion case. Indeed, with helps of this formula we will be able to establish explicit derivative formulae and Harnack inequalities for a class of jump processes (see Sections 3 and 4).

Before move on, let us introduce some recent results concerning regularity properties of the semigroup associated to the following linear SDE

$$\mathrm{d}X_t = A_t X_t \mathrm{d}t + \sigma_t \mathrm{d}L_t \tag{1.1}$$

on $\mathbb{R}^d$, where $A, \sigma : [0, \infty) \to \mathbb{R}^d \otimes \mathbb{R}^d$ are measurable such that σ_s is invertible for $s \geq 0$ and A, σ, σ^{-1} are locally bounded, L_t is the Lévy process on $\mathbb{R}^d$ with Lévy measure ν (see e.g. [1, 13]). Let P_t be the semigroup associated to (1.1), i.e.

$$P_t f(x) = \mathbb{E} f(X_t^x), \quad t \geq 0, x \in \mathbb{R}^d, f \in \mathscr{B}_b(\mathbb{R}^d),$$

where $\mathscr{B}_b(\mathbb{R}^d)$ is the set of all bounded measurable functions on $\mathbb{R}^d$, and X_t^x is the solution with initial data x. To formulate the solution, for any $s \geq 0$ let $(T_{s,t})_{t \geq s}$ solve the equation on $\mathbb{R}^d \otimes \mathbb{R}^d$:

$$\frac{\mathrm{d}}{\mathrm{d}t} T_{s,t} = A_t T_{s,t}, \ T_{s,s} = I.$$

Let $T_t = T_{0,t}, t \geq 0$. We have $T_t = T_{s,t} T_s$ for $t \geq s \geq 0$ and

$$X_t = T_t x + \int_0^t T_{s,t} \sigma_s \mathrm{d}L_s, \quad t \geq 0. \tag{1.2}$$

By using lower bound conditions of the Lévy measure ν, the coupling property and gradient estimates have been derived in [34, 35, 8, 26, 27]. Moreover, by using subordinations, the dimension-free Harnack inequality has been established in [11] for some jump processes in terms of known inequalities in the diffusion setting, where for the log-Harnack inequality the associated Bernstein function can be very broad but for the Harnack inequality with a power the function was assumed to have a growth stronger than $\sqrt{r}$. When A_t and σ_t

are independent of t and $\nu(\mathrm{d}z) \geq c|z|^{-(\mathrm{d}+\alpha)}\mathrm{d}z$ for some constants $c > 0$ and $\alpha \in (0, 2)$, i.e. the equation is time-homogenous with noise having an α-stable part, a different version of Harnack inequality was presented in [39, Theorem 1.1 and Corollary 1.3]: for any $p \geq 1$ there exists a constant $C > 0$ such that

$$(1.3) \qquad (P_t f(x+h))^p \leq C P_t f^p(x) \Big(1 + \frac{|h|}{(t \wedge 1)^{\frac{1}{\alpha}}}\Big)^{p(d+\alpha)}, \quad t > 0, x, h \in \mathbb{R}^d$$

holds for positive $f \in \mathscr{B}_b(\mathbb{R}^d)$. Since here p is allowed to be 1, which is impossible even for the Brownian motion, this inequality is somehow stronger than usual ones derived in the diffusion setting. On the other hand, however, the equality does not hold even for $h = 0$, so that this inequality is weaker than those known ones for small distance.

The remainder of the paper is organized as follows. In the next section we present two lemmas, which will be used to establish derivative formulae and Harnack inequalities in Sections 3 and 4 respectively. In particular, a new Girsanov theorem is presented for Lévy processes using the Lévy measure is presented, which is interesting by itself. With concrete lower bounds of the Lévy measure, explicit gradient estimates and Harnack inequalities will also be addressed in Sections 3 and 4, which extend and improve the corresponding known results derived recently in [34, 39, 27], see Corollaries 3.3 and 4.3 for details.

2 Prelimilary

Form now on, let $t > 0$ be fixed, and let

$$\overline{W}_t = \big\{w : [0, t] \to \mathbb{R}^d \text{ is right-continuous with left limits}\big\}$$

be the path space, which is a Polish space under the Skorokhod metric. For any $w \in \overline{W}_t$ and $s \in [0, t]$, let

$$\Delta w_s = w_s - w_{s-}, \quad w_{s-} = \lim_{s' \uparrow s} w_{s'}.$$

Let $L = (L_s)_{s\in[0,t]}$ be the Lévy process with Lévy measure ν. Then its distribution Λ is a probability measure on

$$W_t := \big\{w \in \overline{W}_t : \ \{s \in [0, t] : |\Delta w_s| \geq \varepsilon\} \text{ is finite for any } \varepsilon > 0\big\}.$$

For any $w \in W_t$,

$$(2.1) \qquad w(\cdot) := \sum_{s\in[0,t],\Delta w_s \neq 0} \delta_{(z,s)}$$

be a σ-finite $\mathbb{Z}_+ \cap \{\infty\}$-valued measure on $\mathbb{R}^d \times [0, t]$. For any non-negative function g on $\mathbb{R}^d \times [0, t]$, let

$$w(g) = \int_{\mathbb{R}^d\times[0,t]} g(z, s) w(\mathrm{d}z, \mathrm{d}s) = \sum_{s\in[0,t],\Delta w_s\neq 0} g(z, s).$$

Now, let $\nu \geq \nu_0$, where ν_0 is another Lévy measure. We may write $L = L^1 + L^0$, where L^1 and L^0 are two independent Lévy processes with Lévy measure $\nu - \nu_0$ and ν_0 respectively, and Λ^0 does not have Gaussian term. L et Λ^1 and Λ^0 be the distributions of L^1 and L^0 respectively. We have $\Lambda = \Lambda^1 * \Lambda^0$. In the sequel we will mainly use the L^0 part to establish derivative formulae of P_t.

It is well known that Λ^0 can be represented by using the Poisson measure Π_t with intensity

$$\mu_t(\mathrm{d}z, \mathrm{d}s) = 1_{[0,t]}(s)\nu_0(\mathrm{d}z) \times \mathrm{d}s,$$

which is a probability measure on the configuration space

$$\begin{aligned}\Gamma_t := \Big\{\gamma := \sum_{i=1}^{n} \delta_{(z_i,s_i)} : n \in \mathbb{Z}_+ \cup \{\infty\}, z_i \in \hat{\mathbb{R}}^d, s_i \in [0,t], \\ \gamma(\{|z| \geq \varepsilon\} \times [0,t]) < \infty \text{ for } \varepsilon > 0\Big\}\end{aligned}$$

equipped with the vague topology, where $\hat{\mathbb{R}}^d = \mathbb{R}^d \setminus \{0\}$. More precisely (see e.g. [14]),

$$\Lambda^0 = \Pi_t \circ \phi^{-1} \tag{2.2}$$

holds for

$$\phi(\gamma) := Bt + \int_{[0,t]\times\{|z|>1\}} z1_{[s,t]}\gamma(\mathrm{d}s, \mathrm{d}z) + \int_{[0,t]\times\{0<|z|\leq 1\}} z1_{[s,t]}(\gamma - \mu_t)(\mathrm{d}s, \mathrm{d}z), \quad \gamma \in \Gamma_t,$$

where $B \in \mathbb{R}^d$ is a constant. Since μ_t is a Lévy measure on $[0,t] \times \mathbb{R}^d$, $\phi(\gamma) \in W_t$ is well-defined for Π_t-a.s. γ. It is easy to see that

$$\phi(\gamma - \delta(z,s)) = \phi(\gamma) - z1_{[s,t]}, \quad \text{for } \delta_{(z,s)} \leq \gamma, \tag{2.3}$$

and by (2.1),

$$\gamma(\mathrm{d}z, \mathrm{d}s) = \phi(\gamma)(\mathrm{d}z, \mathrm{d}s). \tag{2.4}$$

This and the Mecke formula for the Poisson measure imply the following lemma, which is crucial for our study.

Lemma 2.1. *For any $h \in L^1(W_t \times \mathbb{R}^d \times [0,t]; \Lambda^0 \times \nu_0 \times \mathrm{d}s)$,*

$$\begin{aligned}&\int_{W_t\times\mathbb{R}^d\times[0,t]} h(w,z,s)\Lambda^0(\mathrm{d}w)\mu_t(\mathrm{d}z,\mathrm{d}s) \\ &= \int_{W_t} \Lambda^0(\mathrm{d}w) \int_{\mathbb{R}^d\times[0,t]} h(w - z1_{[s,t]}, z, s)w(\mathrm{d}z, \mathrm{d}s).\end{aligned} \tag{2.5}$$

Consequently, for X_t^x solving (1.1) with initial data x,

$$\begin{aligned}&\mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X_t^x + T_{s,t}\sigma_s z)h(L^0, z, s)\mu_t(\mathrm{d}z, \mathrm{d}s) \\ &= \mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X_t^x)h(L^0 - z1_{[s,t]}, z, s)L^0(\mathrm{d}z, \mathrm{d}s).\end{aligned} \tag{2.6}$$

Proof. Combining (2.2) with the Mecke formula [19] (see also [23]), and using (2.3) and (2.4), we obtain

$$\begin{aligned}
&\int_{W_t\times\mathbb{R}^d\times[0,t]} h(w,z,s)\Lambda^0(\mathrm{d}w)\mu_t(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{\Gamma_t}\Pi_t(\mathrm{d}\gamma)\int_{\mathbb{R}^d\times[0,t]} h(\phi(\gamma),z,s)\mu_t(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{\Gamma_t}\Pi_t(\mathrm{d}\gamma)\int_{\mathbb{R}^d\times[0,t]} h(\phi(\gamma-\delta_{(z,s)}),z,s)\gamma(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{\Gamma_t}\Pi_t(\mathrm{d}\gamma)\int_{\mathbb{R}^d\times[0,t]} h(\phi(\gamma)-z1_{[s,t]}),z,s)\phi(\gamma)(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{W_t}\Lambda^0(\mathrm{d}w)\int_{\mathbb{R}^d\times[0,t]} h(w-z1_{[s,t]},z,s)w(\mathrm{d}z,\mathrm{d}s).
\end{aligned}$$

Hence, (2.5) holds.

Next, let

$$\psi(w)=T_tx+\int_0^t T_{s,t}\sigma_s\mathrm{d}w_s,$$

where the integral w.r.t. $\mathrm{d}w_s$ is the Itô integral which is Λ-a.s. defined on W_t. By (1.2) we have

$$f(X_t^x)=f\circ\psi(L^1+L^0),\quad f(X_t^x+T_{s,t}\sigma_s z)=f\circ\psi(L^1+L^0+z1_{[s,t]}).$$

Combining this with (2.5) and noting that L^1 and L^0 are independent with distributions Λ^1 and Λ^0 respectively, we obtain

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X_t^x+T_{s,t}\sigma_s z)h(L^0,z,s)\mu_t(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]} f\circ\psi(w^1+w^0+z1_{[s,t]})h(w^0,z,s)\Lambda^0(\mathrm{d}w^0)\mu_t(\mathrm{d}z,\mathrm{d}s)\\
&=\int_{W_t\times W_t}\Lambda^1(\mathrm{d}w^1)\Lambda^0(\mathrm{d}w^0)\int_{\mathbb{R}^d\times[0,t]} f\circ\psi(w^1+w^0)h(w^0-z1_{[s,t]},z,s)w^0(\mathrm{d}z,\mathrm{d}s)\\
&=\mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X_t^x)h(L^0-z1_{[s,t]},z,s)L^0(\mathrm{d}z,\mathrm{d}s).
\end{aligned}$$

□

As an application of Lemma 2.1, we have the following Girsanov theorem, which might be interesting by itself.

Theorem 2.2. *Let $G\ge 0$ be a measurable function on $W_t\times\mathbb{R}^d\times[0,t]$ such that $(\Lambda^0\times\mu_t)(G)=1$. Let (ξ,τ) be a random variable on $\mathbb{R}^d\times[0,t]$ such that the distribution of (L^0,ξ,τ) is $G(w,z,s)\Lambda^0(\mathrm{d}w)\mu_t(\mathrm{d}z)$. Let*

$$g(z,s)=\int_{W_t} G(w,z,s)\Lambda^0(\mathrm{d}w),$$

which is the distribution density of (ξ, τ) w.r.t. μ_t. If $\mu_t(g > 0) = \infty$ and

$$1_{\{G(w,z,s)>0\}}g(z, s) = g(z, s), \quad (\Lambda^0 \times \mu_t)\text{−a.e.}$$

Then the process

$$L^0 + \xi 1_{[\tau,t]} := \big(L_s^0 + \xi 1_{[\tau,t]}(s)\big)_{s\in[0,t]}$$

has distribution Λ^0 under the probability measure $\mathbb{Q} := R\mathbb{P}$, where

$$R = \frac{g(\xi, \tau)}{G(L^0, \xi, \tau)\{L^0(g) + g(\xi, \tau)\}}.$$

Proof. Since G is the distribution density of (L^0, ξ, τ) w.r.t. $\Lambda^0 \times \mu_t$, we have $G(L^0, \xi, \tau) > 0$ a.s. Similarly, $g(\xi, \tau) > 0$ a.s. as well. Moreover, it is easy to see that $\mathbb{E}L^0(g) = 1$ so that $w(g) < \infty$. Therefore, $R \in (0, \infty)$ a.s.

Now, for any non-negative measurable function F on W_t, applying (2.5) to

$$h(w, z, s) = \frac{F(w, +z1_{[s,t]})}{g(z, s) + w(g)} 1_{\{G>0\}}(w, z, s),$$

which is finite since $\mu_t(g > 0) = \infty$ implies that $w(g) > 0$ holds Λ^0-a.e., and using $1_{\{G(w,z,s)>0\}}g(z, s) = g(z, s)$, we obtain

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}F(L^0 + \xi 1_{[\tau,t]}) &= \mathbb{E}\frac{F(L^0 + \xi 1_{[\tau,t]})g(\xi, \tau)}{G(L^0, \xi, \tau)\{g(\xi, \tau) + L^0(g)\}} \\
&= \int_{G(w,z,s)>0} \frac{F(w + z1_{[s,t]})G(w, z, s)g(z, s)}{G(w, z, s)\{g(z, s) + w(g)\}}\Lambda^0(\mathrm{d}w)\mu_t(\mathrm{d}z, \mathrm{d}s) \\
&= \int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{F(w + z1_{[s,t]})g(z, s)}{g(z, s) + w(g)}\Lambda^0(\mathrm{d}w)\mu_t(\mathrm{d}z, \mathrm{d}s) \\
&= \int_{W_t} \Lambda^0(\mathrm{d}w) \int_{\mathbb{R}^d\times[0,t]} \frac{F(w)g(z, s)}{w(g)} w(\mathrm{d}z, \mathrm{d}s) = \int_{W_t} F(w)\Lambda^0(\mathrm{d}w).
\end{aligned}$$

This completes the proof. □

A simple choice of G in the above Theorem is that $G(w, z, s) = g(z, s)$, i.e. (ξ, τ) is independent of L^0. To derive gradient estimate from Theorem 3.1 below, we need the following Γ-function:

$$\Gamma(r) = \int_0^\infty s^{r-1}\mathrm{e}^{-rs}\mathrm{d}s, \quad r > 0.$$

Lemma 2.3. *Let Λ^0 be the distribution of a Lévy process with Lévy measure ν_0 which is not necessarily absolutely continuous w.r.t. the Lebesgue measure. Let $\mu_t(\mathrm{d}z, \mathrm{d}s) = \nu_0(\mathrm{d}z) \times \mathrm{d}s$ on $\mathbb{R}^d \times [0, t]$. Then for any non-negative measurable function g on $\mathbb{R}^d \times [0, t]$,*

$$\int_{W_t} \frac{\Lambda^0(\mathrm{d}w)}{w(g)^\theta} = \frac{1}{\Gamma(\theta)} \int_0^\infty r^{\theta-1} \exp\Big[- \mu_t(1 - \mathrm{e}^{-rg})\Big]\mathrm{d}r, \quad \theta > 0.$$

Proof. Noting that

$$\frac{1}{s^\theta} = \frac{1}{\Gamma(\theta)} \int_0^\infty r^{\theta-1} \mathrm{e}^{-sr} \mathrm{d}r, \quad s > 0,$$

it follows from (2.2) that

$$\begin{aligned}\int_{W_t} \frac{\Lambda^0(\mathrm{d}w)}{w(g)^\theta} &= \int_{\Gamma_t} \frac{\Pi_t(\mathrm{d}\gamma)}{\gamma(g)^\theta} = \frac{1}{\Gamma(\theta)} \int_0^\infty r^{\theta-1} \mathrm{d}r \int_{\Gamma_t} \mathrm{e}^{-r\gamma(g)} \Pi_t(\mathrm{d}\gamma)\\ &= \frac{1}{\Gamma(\theta)} \int_0^\infty r^{\theta-1} \exp\Big[-\mu_t(1-\mathrm{e}^{-rg})\Big] \mathrm{d}r.\end{aligned}$$

□

3 Derivative formula and gradient estimates

To establish a derivative formula for P_t, we need an absolutely continuous lower bound of ν. Let

$$\nu(\mathrm{d}z) \geq \nu_0(\mathrm{d}z) := \rho_0(z)\mathrm{d}z$$

such that $\nu_0(\mathbb{R}^d) = \infty$. Recall that the infinity of ν is essential to ensure the strong Feller property of P_t, which is necessary for the differentiability of the semigroup (see [22] and references within for criteria on the strong Feller property). Thus, the assumption $\nu_0(\mathbb{R}^d) = \infty$ is reasonable in order to establish a derivative formula of P_t.

Let $L^0 = (L^0_s)_{s\in[0,t]}$ be the Lévy process with Lévy measure ν_0, and let $L^1 = (L^1_s)_{s\in[0,t]}$ be the Lévy processes with Lévy measures $\nu_1 := \nu - \nu_0$ independent of L^0, so that $L := L^0 + L^1$ is the Lévy process with Lévy measure ν introduced above. Let $\hat{\mathbb{R}}^d = \mathbb{R}^d \setminus \{0\}$ and let $\nu_0(g)$ be the integral of g w.r.t. ν_0.

Theorem 3.1. *Let $\nu(\mathrm{d}z) \geq \rho_0(z)\mathrm{d}z$ for some non-negative $\rho_0 \in W^{1,1}_{loc}(\hat{\mathbb{R}}^d)$ such that $\nu_0(\mathrm{d}z) := \rho_0(z)\mathrm{d}z$ is an infinite measure. Let Λ^0 be the distribution of L^0 given above. If there exists a non-negative measurable function g on $\mathbb{R}^d \times [0,t]$ differentiable in $z \in \mathbb{R}^d$ such that*

$$\int_0^\infty \exp\Big[-\mu_t(1-\mathrm{e}^{-rg})\Big]\mathrm{d}r + \int_{\mathbb{R}^d\times[0,t]} \big\{\rho_0 g + \rho_0|\nabla g| + g|\nabla\rho_0|\big\}(z,s)\mathrm{d}z\mathrm{d}s < \infty, \tag{3.1}$$

where ∇ is the gradient in $z \in \mathbb{R}^d$, then for any $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\begin{aligned}&\nabla P_t f(x)\\ &= -\mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X^x_t + T_{s,t}\sigma_s z)\frac{(\sigma_s^{-1}T_s)^*\{L^0(g)\nabla(\rho_0 g) + g^2\nabla\rho_0\}}{(L^0(g)+g)^2}(z,s)\mathrm{d}z\mathrm{d}s\\ &= \mathbb{E}\bigg[f(X^x_t)\int_{\mathbb{R}^d\times[0,t]} \frac{(\sigma_s^{-1}T_s)^*\{\rho_0 g\nabla g - L^0(g)\nabla(\rho_0 g)\}}{L^0(g)^2\rho_0}(z,s)L^0(\mathrm{d}z,\mathrm{d}s)\bigg].\end{aligned} \tag{3.2}$$

Proof. Noting that the second equality in (3.2) follows from the first and (2.6), we only need to prove the first formula.

(a) We first prove for the case where g has a compact support K. Let $\Lambda = \Lambda^0 * \Lambda^1$ be the distribution of L. For $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $\varepsilon \in (0,1)$, let

$$h_\varepsilon(w) = f\bigg(T_t(x+\varepsilon h) + \int_0^t T_{s,t}\sigma_s \mathrm{d}w_s\bigg).$$

By (1.2) and noting that L^0 and L^1 are independent with distributions Λ^0 and Λ^1 respectively, we have $f(X_t^{x+\varepsilon h}) = h_\varepsilon(L^0 + L^1)$ and

$$P_t f(x+\varepsilon h) = \int_{W_t\times W_t} h_\varepsilon(w^1 + w^0)\Lambda^0(\mathrm{d}w^0)\Lambda^1(\mathrm{d}w^1).$$

Since $T_t = T_{s,t}T_s$ for $s \in [0,t]$, and since due to (3.1) and Lemma 2.3 $w(g) > 0$ holds for Λ^0-a.s. $w \in W_t$, this implies

$$\begin{aligned}P_t f(x+\varepsilon h) &= \int_{W_t\times W_t} \Lambda^1(\mathrm{d}w^1)\Lambda^0(\mathrm{d}w^0)\int_{\mathbb{R}^d\times[0,t]} \frac{h_\varepsilon(w^1+w^0)g(z,s)}{w^0(g)} w^0(\mathrm{d}z,\mathrm{d}s)\\ &= \int_{W_t\times W_t} \Lambda^1(\mathrm{d}w^1)\Lambda^0(\mathrm{d}w^0)\int_{\mathbb{R}^d\times[0,t]} \frac{h_0(w^1+w^0+\varepsilon\sigma_s^{-1}T_s h 1_{[s,t]})g(z,s)}{w^0(g)} w^0(\mathrm{d}z,\mathrm{d}s).\end{aligned}$$

Combining this with (2.5) for Λ^0 in place of Λ and

$$h(w,z,s) := \frac{h_0(w^1 + w + (z+\varepsilon\sigma_s^{-1}T_s h)1_{[s,t]})g(z,s)}{(w^0 + z1_{[s,t]})(g)}, \quad w \in W_t,$$

we arrive at

$$\begin{aligned}&P_t f(x+\varepsilon h)\\ &= \int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{h_0(w^1+w^0+(z+\varepsilon\sigma_s^{-1}T_s h)1_{[s,t]})g(z,s)}{w^0(g)+g(z,s)}\Lambda^0(\mathrm{d}w^0)\mu_t(\mathrm{d}z,\mathrm{d}s)\\ &= \int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{h_0(w^1+w^0+(z+\varepsilon\sigma_s^{-1}T_s h)1_{[s,t]})(\rho_0 g)(z,s)}{w^0(g)+g(z,s)}\Lambda^0(\mathrm{d}w^0)\mathrm{d}z\mathrm{d}s.\end{aligned}$$

Using the integral transform $z \mapsto z - \varepsilon\sigma_s^{-1}T_s h$, it follows that

$$\begin{aligned}&P_t f(x+\varepsilon h) =\\ &\int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{h_0(w^1+w^0+z1_{[s,t]})(\rho_0 g)(z-\varepsilon\sigma_s^{-1}T_s h,s)}{w^0(g)+g(z-\varepsilon\sigma_s^{-1}T_s h,s)}\Lambda^0(\mathrm{d}w^0)\mathrm{d}z\mathrm{d}s.\end{aligned}\tag{3.3}$$

Therefore,

$$\begin{aligned}&\frac{P_t f(x+\varepsilon h) - P_t f(x)}{\varepsilon}\\ &= \int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]} h_0(w^1+w^0+z1_{[s,t]})\Phi_\varepsilon(w^0,z,s)\Lambda^0(\mathrm{d}w^0)\mathrm{d}z\mathrm{d}s\end{aligned}\tag{3.4}$$

holds for

$$\Phi_\varepsilon(w^0,z,s):=\frac{1}{\varepsilon}\bigg(\frac{g\rho_0}{w^0(g)+g}(z-\varepsilon\sigma_s^{-1}T_sh,s)-\frac{g\rho_0}{w^0(g)+g}(z,s)\bigg).$$

Since

$$\begin{aligned}\lim_{\varepsilon\to0}\Phi_\varepsilon(w,z,s)&=-\Big\langle\nabla\frac{\rho_0 g}{w(g)+g}(z,s),\sigma_s^{-1}T_sh\Big\rangle\\&=-\Big\langle(\sigma_s^{-1}T_s)^*\frac{w(g)\nabla(\rho_0 g)+g^2\nabla\rho_0}{(w(g)+g)^2}(z,s),h\Big\rangle,\end{aligned}$$

to derive the desired derivative formula by letting $\varepsilon\to0$, we need to make use of the dominated convergence theorem. Since g has a compact support and $\sup_{s\in[0,t]}|\sigma_s^{-1}T_sh|<\infty$, there is a compact set $K\subset\mathbb{R}^d$ such that $\operatorname{supp}\Phi_\varepsilon\subset W_t\times K\times[0,t]$ holds for all $\varepsilon\in(0,1)$. Since $\Lambda^0(\mathrm{d}w)\times\mathrm{d}z\times\mathrm{d}s$ is finite on $W_t\times K\times[0,t]$, it suffices to show that $\{\Phi_\varepsilon\}_{\varepsilon\in(0,1)}$ is uniformly integrable w.r.t. this measure. Noting that

$$\Big|\nabla\frac{\rho_0 g}{w(g)+g}\Big|\le\frac{|\nabla(\rho_0 g)|}{w(g)+g}+\frac{\rho_0 g|\nabla g|}{(w(g)+g)^2}\le\frac{2(\rho_0|\nabla g|+g|\rho_0|)}{w(g)},$$

there exists a constant $c>0$ such that we have

$$\begin{aligned}|\Phi_\varepsilon(w^0,z,s)|&=\bigg|\frac{1}{\varepsilon}\int_0^\varepsilon\Big\{\frac{\mathrm{d}}{\mathrm{d}r}\Big(\frac{\rho_0 g}{w^0(g)+g}\Big)(z-r\sigma_s^{-1}T_sh,s)\Big\}\mathrm{d}r\bigg|\\&\le\frac{c}{\varepsilon w(g)}\int_0^\varepsilon\big\{\rho_0|\nabla g|+g|\nabla\rho_0|\big\}(z-r\sigma_s^{-1}T_sh,s)\mathrm{d}r.\end{aligned}$$

By (3.1) and Lemma 2.3 we see that $\int_{W_t}\frac{1}{w(g)}\Lambda^0(\mathrm{d}w)<\infty$. So, it suffices to show that

$$\Psi_\varepsilon(z,s):=\frac{1}{\varepsilon}\int_0^\varepsilon\big\{\rho_0|\nabla g|+g|\nabla\rho_0|\big\}(z-r\sigma_s^{-1}T_sh,s)\mathrm{d}r,\quad\varepsilon\in(0,1)$$

is uniformly integrable w.r.t. $\mathrm{d}z\times\mathrm{d}s$ on $K\times[0,t]$. Since the function $s\mapsto(s-R)^+$ is convex, by the Jensen inequality we have

$$(\Psi_\varepsilon-R)^+(z,s)\le\frac{1}{\varepsilon}\int_0^\varepsilon\big(\rho_0|\nabla g|+g|\nabla\rho_0|-R\big)^+(z-r\sigma_s^{-1}T_sh,s)\mathrm{d}r.$$

So,

$$\begin{aligned}&\int_{K\times[0,t]}(\Psi_\varepsilon(z,s)-R)^+\mathrm{d}z\mathrm{d}s\\&\le\frac{1}{\varepsilon}\int_{\mathbb{R}^d\times[0,t]\times[0,\varepsilon]}\big(\rho_0|\nabla g|+g|\nabla\rho_0|-R\big)^+(z-r\sigma_s^{-1}T_sh,s)\mathrm{d}z\mathrm{d}s\mathrm{d}r\\&=\int_{\mathbb{R}^d\times[0,t]}\big(\rho_0|\nabla g|+g|\nabla\rho_0|-R\big)^+(z,s)\mathrm{d}z\mathrm{d}s,\quad\varepsilon\in(0,1),\end{aligned}$$

where the last step is due to the integral transform $z \mapsto z + r\mathrm{e}^{-sA}h$ for the integral w.r.t. $\mathrm{d}z$. Combining this with (3.1) we see that

$$\lim_{R\to\infty} \sup_{\varepsilon\in(0,1)} \int_{K\times[0,t]} (\Psi_\varepsilon(z,s) - R)^+ \mathrm{d}z\mathrm{d}s = 0,$$

that is, $\{\Psi_\varepsilon\}_{\varepsilon\in(0,1)}$ is uniform integrable w.r.t. $\mathrm{d}z \times \mathrm{d}s$ on $K \times [0,t]$.

(b) Let g satisfy (3.1). For any $n \geq 1$, let $g_n(z,s) = g(z,s)\{1 \wedge (1+n-|z|)^+\}$ which has a compact support. By (a) we have

$$\nabla P_t f(x) = -\mathbb{E} \int_{\mathbb{R}^d\times[0,t]} f(X_t^x + \mathrm{e}^{(t-s)A}z) \frac{(\sigma_s^{-1}T_s)^*\{L^0(g_n)\nabla(g_n\rho_0) + g_n^2\nabla\rho_0\}}{(L^0(g_n)+g_n)^2}(z,s)\mathrm{d}z\mathrm{d}s \tag{3.5}$$

Let $c = \mathrm{e}^{t\|A\|}$. It is easy to see that

$$\left|\frac{(\sigma_s^{-1}T_s)^*\{L^0(g_n)\nabla(g_n\rho_0) + g_n^2\nabla\rho_0\}}{(L^0(g_n)+g_n)^2}\right| \leq \frac{c\{\rho_0|\nabla g| + g|\nabla\rho_0| + \rho_0 g\}}{L^0(g_1)}, \quad n \geq 1$$

holds for some constant $c > 0$. Then, according to (3.1), the desired formula follows from the dominated convergence theorem by letting $n \to \infty$ in (3.5), provided

$$\int_{W_t} \frac{\Lambda^0(\mathrm{d}w)}{w(g_1)} < \infty. \tag{3.6}$$

By Lemma 2.3 and (3.1) we have

$$\begin{aligned}\int_{W_t} \frac{\Lambda^0(\mathrm{d}w)}{w(g_1)} &= \frac{1}{\Gamma(1)} \int_0^\infty \exp\Big[-\mu_t(1-\mathrm{e}^{-rg_1})\Big]\mathrm{d}r\\ &\leq \frac{1}{\Gamma(1)} \int_0^\infty \exp\Big[t\nu_0(|z|>1) - \mu_t(1-\mathrm{e}^{-rg})\Big]\mathrm{d}r < \infty\end{aligned}$$

since $\nu_0(|z| \geq 1) \leq \nu(|z| \geq 1) < \infty$ as ν is a Lévy measure. Therefore, the proof is finished. □

Corollary 3.2. *Let $\rho_0 \in W^{1,1}_{loc}(\hat{\mathbb{R}}^d)$ be non-negative such that $\nu(\mathrm{d}z) \geq \rho_0(z)\mathrm{d}z$, and let g be a non-negative measurable function on $\mathbb{R}^d \times [0,t]$ differentiable in the first variable such that $\mu_t(g) < \infty$. Then for any $p \in (1,\infty]$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$,*

$$\begin{aligned}|\nabla P_t f| \leq &(P_t|f|^p)^{\frac{1}{p}} \bigg(\frac{1}{\Gamma(1)}\int_0^\infty \exp\Big[-\mu_t(1-\mathrm{e}^{-rg})\Big]\mathrm{d}r\bigg)^{\frac{p-1}{p}}\\ &\times \bigg(\int_{\mathbb{R}^d\times[0,t]} \Big\{\|\sigma_s^{-1}T_s\|\big(g|\nabla\log\rho_0| + |\nabla\log(\rho_0 g)|\big)(z,s)\Big\}^{\frac{p}{p-1}} g(z,s)\mu_t(\mathrm{d}z,\mathrm{d}s)\bigg)^{\frac{p-1}{p}}.\end{aligned}$$

Proof. Assume that the desired upper bound is finite. Then (3.1) holds. On the other hand, according to (2.6), for any $f \in \mathscr{B}_b(\mathbb{R}^d)$ we have

$$P_t f(x) = \mathbb{E}\int_{W_t\times\mathbb{R}^d\times[0,t]} f(X_t^x + T_{s,t}\sigma_s z)\frac{(g\rho_0)(z,s)}{L^0(g)+g(z,s)}\mathrm{d}z\mathrm{d}s. \tag{3.7}$$

Combining this with the first formula in (3.2) and using the Hölder inequality, we obtain

$$\begin{aligned}&|\nabla P_t f|\\ &\le \mathbb{E}\int_{\mathbb{R}^d\times[0,t]} f(X_t^x + T_{s,t}\sigma_s z)\|\sigma_s^{-1}T_s\|\Big\{\big(g|\nabla\log\rho_0| + |\nabla\log(g\rho_0)|\big)\frac{(g\rho_0)}{L^0(g)+g}\Big\}(z,s)\mathrm{d}z\mathrm{d}s\\ &\le (P_t|f|^p)^{\frac{1}{p}}\bigg(\mathbb{E}\int_{\mathbb{R}^d\times[0,t]}\Big\{\|\sigma_s^{-1}T_s\|\big(g|\nabla\log\rho_0| + |\nabla\log(\rho_0 g)|\big)\Big\}^{\frac{p}{p-1}}\frac{g}{L^0(g)+g}\mathrm{d}\mu_t\bigg)^{\frac{p-1}{p}}.\end{aligned}$$

Then the desired gradient estimate follows from Lemma 2.3 by noting that $\frac{g}{L^0(g)+g} \le \frac{g}{L^0(g)}$ and $\mathbb{E}\frac{1}{L^0(g)} = \int_{W_t}\frac{1}{w(g)}\Lambda^0(\mathrm{d}w)$. $\square$

To illustrate Corollary 3.2, we present explicit conditions on the lower bound of ν for the gradient estimate and Harnack inequality. Comparing with results in [34, 27] where the uniform gradient estimates are derived, in the following result ρ_0 is not necessary corresponding to a Bernstein function and more general L^p gradient estimates are also provided. For a $d\times d$ matrix M and a constant $\alpha\in\mathbb{R}$, we write $M \le \alpha I$ provided $\langle Ma, a\rangle \le \alpha|a|^2$ holds for all $a \in \mathbb{R}^d$.

Corollary 3.3. *Let $A_s \le \alpha I$ and $\|\sigma_s^{-1}\| \le \lambda$ for some constants $\alpha\in\mathbb{R}$ and $\lambda > 0$. Let $\nu(\mathrm{d}z) \ge |z|^{-d}S(|z|^{-2})1_{\{s\le r_0\}}$ for some constant $r_0 > 0$ and positive function $S \in C^1([0,\infty))$ such that*

$$\limsup_{r\to\infty}\frac{|S'(r)|r}{S(r)} < \infty. \tag{3.8}$$

For $p > 1$ and $k > 2 + \frac{p}{p-1}$, let

$$\psi_k(r) = \frac{(1-\mathrm{e}^{-1})\kappa(d)}{2^k}\int_{\frac{r_0}{2}\wedge r^{-1/k}}^{\frac{r_0}{2}}\frac{S(s^{-2})}{s}\mathrm{d}s,\quad r > 0,$$

where $\kappa(d)$ is the area of the unit sphere in $\mathbb{R}^d$. If $\int_0^\infty \mathrm{e}^{-t\psi_k(r)}\mathrm{d}r < \infty$, then there exists a constant $c > 0$ such that

$$|\nabla P_t f| \le c\Big(\frac{\mathrm{e}^{\alpha t}-1}{\alpha}\Big)^{\frac{p-1}{p}}(P_t|f|^p)^{\frac{1}{p}}\bigg(\int_0^\infty \mathrm{e}^{-t\psi_k(r)}\mathrm{d}r\bigg)^{\frac{p-1}{p}}$$

holds for $f\in\mathscr{B}_b(\mathbb{R}^d)$, In particular, if $S(r) = c_0\log^\varepsilon(1+r)$ for some $c_0, \varepsilon > 0$, then for any $p > 1$ there exists a constant $c > 0$ such that

$$|\nabla P_t f| \le (P_t f^p)^{1/p}\exp\big[c(1\wedge t)^{-\frac{1}{\varepsilon}}\big],\quad t > 0$$

holds for all positive $f\in\mathscr{B}_b(\mathbb{R}^d)$.

Proof. Let $\rho_0(z) = |z|^{-d}S(|z|^{-2})(1-\frac{|z|}{r_0})^{+k}$ and $g(z,s) = g(z) = |z|^k$. Obviously, $\int_0^\infty \mathrm{e}^{-t\psi_k(r)}\mathrm{d}r < \infty$ implies that $\int_{\mathbb{R}^d}\rho_0(z)\mathrm{d}z = \infty$. Since $\|S'\|_\infty < \infty$ implies that $S(s^{-2}) \le cs^{-2}$ for some constant $c > 0$ and all $s \le r_0$, we have

$$(g\rho_0)(z) \le c|z|^{k-d-2}1_{\{|z|\le r_0\}},$$

so that $\mu_t(g) = t\int_{\mathbb{R}^d}(\rho_0 g)(z)\mathrm{d}z < \infty$. Next, it is easy to see from (3.8) that

$$\big\{g|\nabla\log\rho_0| + |\nabla\log(\rho_0 g)|\big\}(z) \le \frac{c}{|z|(r_0-|z|)^+},\quad |z| < r_0$$

holds for some constant $c > 0$. Thus, there exists a constant $c > 0$ such that

$$\Big\{\rho_0 g\big(|g\nabla\log\rho_0|^{\frac{p}{p-1}} + |\nabla\log(\rho_0 g)|^{\frac{p}{p-1}}\big)\Big\}(z) \le c|z|^{k-d-2-\frac{p}{p-1}}(r_0-|z|)^{+k-\frac{p}{p-1}}$$

holds for some constant $c > 0$. Since $\|\sigma_s^{-1}T_s\| \le \lambda\mathrm{e}^{\alpha s}$ and $k > 2+\frac{p}{p-1}$, this implies

$$\begin{aligned}
&\int_{\mathbb{R}^d\times[0,t]}\Big\{\|\sigma_s^{-1}T_s\|\big(g|\nabla\log\rho_0| + |\nabla\log(\rho_0 g)|\big)(z,s)\Big\}^{\frac{p}{p-1}} g(z,s)\mu_t(\mathrm{d}z,\mathrm{d}s)\\
&\le \frac{\lambda(\mathrm{e}^{\alpha t}-1)}{\alpha}\int_{\mathbb{R}^d}\Big\{\rho_0 g\big(|g\nabla\log\rho_0|^{\frac{p}{p-1}} + |\nabla\log(\rho_0 g)|^{\frac{p}{p-1}}\big)\Big\}(z)\mathrm{d}z \le \frac{c(\mathrm{e}^{\alpha t}-1)}{\alpha}
\end{aligned} \tag{3.9}$$

for some constant $c > 0$. Next, for $r \ge (2/r_0)^k$ we have

$$\begin{aligned}
\nu_0(1-\mathrm{e}^{-rg}) &\ge \frac{1}{2^k}\int_{|z|\le r_0/2}|z|^{-d}S(|z|^{-2})(1-\mathrm{e}^{-r|z|^k})\mathrm{d}z\\
&= \frac{\kappa(d)}{2^k}\int_0^{r_0/2}s^{-1}S(s^{-2})(1-\mathrm{e}^{-rs^k})\mathrm{d}s\\
&\ge \frac{\kappa(d)}{2^k}\int_{\frac{r_0}{2^k}\wedge r^{-1/k}}^{\frac{r_0}{2}}\frac{S(s^{-2})(1-\mathrm{e}^{-1})}{s}\mathrm{d}s = \psi_k(r).
\end{aligned}$$

Combining this with (3.9), we prove the first assertion by Corollary 3.2.

Next, let $S(r) = c_0\log^\varepsilon(1+r)$. By the semigroup property and the Jensen inequality, it suffices to prove the desired gradient estimate for $t \in (0,1]$. It is easy to see that

$$t\psi_k(r) \ge c_1 t\log^{1+\varepsilon}(1+r) - c_2 t \ge 2\log(1+r) - c_3 t^{-1/\varepsilon}$$

holds for some constants $c_1, c_2, c_3 > 0$. Then the desired gradient estimate follows from the first part of this Corollary. □

Note that the second estimate in Corollary 3.3 improves and extends [34, Example 1.3] to L^p gradient estimate with better short time behavior. On the other hand, however, Corollary 3.3 does not provide sharp estimate for the α-stable case. In general, to drive sharper gradient estimates, it might be necessary to take g depending also on s.

4 Harnack inequality and applications

We first investigate the Harnack inequality with a power in the sense of [31] and the log-Harnack inequality introduced in [33, 25], then present some applications of these inequalities in an abstract framework. Recently, these type of inequalities have been established in [11] for some jump processes with using subordinations from diffusion processes and in [39] using heat kernel bounds of the α-stable process.

4.1 Harnack inequality

For positive measurable functions ρ_0, g on $\hat{\mathbb{R}}^d$, let $\nu_0(\mathrm{d}z) = \rho_0(z)\mathrm{d}z$ and

$$\gamma_{\rho_0,g}(\theta,t) = \frac{1}{\Gamma(\theta)}\int_0^\infty r^{\theta-1}\exp\big[-t\nu_0(1-\mathrm{e}^{-rg})\big]\mathrm{d}r, \quad \theta, t>0.$$

Theorem 4.1. *Let $\alpha\in\mathbb{R}$ and $\lambda\geq 0$ be such that $A_s\leq\alpha I$ and $\|\sigma_s^{-1}\|\leq\lambda$ for $s\in[0,1]$. Let $\rho_0\in W^{1,1}_{loc}(\hat{\mathbb{R}}^d)$ and $g\in W^{1,1}_{loc}(\mathbb{R}^d)$ be positive such that $\nu(\mathrm{d}z)\geq\rho_0(z)\mathrm{d}z$ and $\nu_0(g>0)=\infty$. Then for any $p>1$ and positive $f\in\mathscr{B}_b(\mathbb{R}^d)$,*

$$\begin{aligned}
&\frac{(P_tf)^p(x+h)}{P_tf^p(x)}\\
&\leq\bigg\{\int_{W_t\times\mathbb{R}^d\times[0,t]}\Big(\frac{(\rho_0g)(z)}{w(g)+g(z)}\Big)^{\frac{p}{p-1}}\Big(\frac{w(g)+g(z+\sigma_s^{-1}T_sh)}{(\rho_0g)(z+\sigma_s^{-1}T_sh)}\Big)^{\frac{1}{p-1}}\Lambda^0(\mathrm{d}w)\mathrm{d}z\mathrm{d}s\bigg\}^{p-1}\\
&\leq\exp\Big[\|\nabla\log(\rho_0g)\|_\infty\lambda\mathrm{e}^\alpha|h|\Big]\bigg\{1+\big(\lambda\|\nabla g\|_\infty\mathrm{e}^\alpha|h|\big)^{\frac{1}{(p-1)\vee 1}}\gamma_{\rho_0,g}\Big(\frac{1}{p-1},t\wedge 1\Big)^{(p-1)\wedge 1}\bigg\}^{(p-1)\vee 1}
\end{aligned}$$

holds for $x,h\in\mathbb{R}^d$ and $t>0$.

Proof. Since $\nu_0(g>0)=\infty$ implies $w(g)>0$ for Λ^0-a.e. w, the right-hand side of the first inequality makes sense (could be infinite). Let $g(z,s)=g(z)$. By (3.3) and the Hörlder inequality, we obtain

$$\begin{aligned}
&(P_tf(x+h))^p\\
&=\bigg\{\int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]}\frac{h_0(w^1+w^0+z1_{[s,t]})(\rho_0g)(z-\sigma_s^{-1}T_sh)}{w^0(g)+g(z-\sigma_s^{-1}T_sh)}\Lambda^0(\mathrm{d}w^0)\mathrm{d}z\mathrm{d}s\bigg\}^p\\
&\leq\bigg\{\int_{W_t}\Lambda^1(\mathrm{d}w^1)\int_{W_t\times\mathbb{R}^d\times[0,t]}\frac{h_0^p(w^1+w^0+z1_{[s,t]})(\rho_0g)(z)}{w^0(g)+g(z)}\Lambda^0(\mathrm{d}w^0)\mathrm{d}z\mathrm{d}s\bigg\}\\
&\qquad\times\bigg\{\int_{W_t\times\mathbb{R}^d\times[0,t]}\Big(\frac{(\rho_0g)(z-\sigma_s^{-1}T_sh)}{w(g)+g(z-\sigma_s^{-1}T_sh)}\Big)^{\frac{p}{p-1}}\Big(\frac{w(g)+g(z)}{(\rho_0g)(z)}\Big)^{\frac{1}{p-1}}\Lambda^0(\mathrm{d}w)\mathrm{d}z\mathrm{d}s\bigg\}^{p-1}\\
&=P_tf^p(x)\bigg\{\int_{W_t\times\mathbb{R}^d\times[0,t]}\Big(\frac{(\rho_0g)(z)}{w(g)+g(z)}\Big)^{\frac{p}{p-1}}\Big(\frac{w(g)+g(z+\sigma_s^{-1}T_sh)}{(\rho_0g)(z+\sigma_s^{-1}T_sh)}\Big)^{\frac{1}{p-1}}\Lambda^0(\mathrm{d}w)\mathrm{d}z\mathrm{d}s\bigg\}^{p-1},
\end{aligned}$$

where in the last step we have used the transform $z \mapsto z + \sigma_s^{-1} T_s h$ for the integral w.r.t. $\mathrm{d}z$. This proves the first inequality.

Next, due to the semigroup property and the Jensen inequality, for the second inequality it suffices to consider $t \in (0.1]$. Then

$$(P_t f(x+h))^p \le P_t f^p(x) \bigg(\int_{W_t \times \mathbb{R}^d \times [0,t]} \frac{g(z)}{w(g)+g(z)} (B_1 B_2)^{\frac{1}{p-1}} \Lambda^0(\mathrm{d}w) \nu_0(\mathrm{d}z) \mathrm{d}s \bigg)^{p-1} \tag{4.1}$$

holds for

$$B_1 = B_1(w,z,s) := \frac{w(g) + g(z + \sigma_s^{-1} T_s h)}{w(g)+g(z)}, \quad B_2 = B_2(w,z,s) := \frac{(\rho_0 g)(z)}{(\rho_0 g)(z + \sigma_s^{-1} T_s h)}.$$

Since $t \in (0,1]$ and $\|\sigma_s^{-1} T_s\| \le \lambda \mathrm{e}^\alpha$ for $s \in (0,1]$, we have

$$\begin{aligned} B_1 \le & 1 + \frac{|g(z+\sigma_s^{-1} T_s h) - g(z)|}{w(g)+g(z)} \le 1 + \frac{\lambda \|\nabla g\|_\infty \mathrm{e}^\alpha |h|}{w(g)+g(z)}, \\ B_2 \le & \exp\Big[\|\nabla \log(\rho_0 g)\|_\infty \lambda \mathrm{e}^\alpha |h| \Big], \quad s \in [0,t]. \end{aligned}$$

Moreover, due to Lemma 2.1

$$\begin{aligned} & \int_{W_t \times \mathbb{R}^d \times [0,t]} \Big(1 + \frac{c}{w(g)+g(z)}\Big)^{\frac{1}{p-1}} \frac{g(z)}{w(g)+g(z)} \Lambda^0(\mathrm{d}w) \nu_0(\mathrm{d}z) \mathrm{d}s \\ &= \int_{W_t} \Big(1 + \frac{c}{w(g)}\Big)^{\frac{1}{p-1}} \Lambda^0(\mathrm{d}w) \int_{\mathbb{R}^d \times [0,t]} \frac{g(z)}{w(g)} \nu_0(\mathrm{d}z) \mathrm{d}s \\ &= \int_{W_t} \Big(1 + \frac{c}{w(g)}\Big)^{\frac{1}{p-1}} \Lambda^0(\mathrm{d}w) \end{aligned} \tag{4.2}$$

holds for $c \ge 0$. So, it follows from (4.1) that

$$\frac{(P_t f(x+h))^p}{P_t f^p(x)} \le \exp\Big[\|\nabla \log(\rho_0 g)\|_\infty \lambda \mathrm{e}^\alpha |h| \Big] \bigg(\int_{W_t} \Big(1 + \frac{\|\nabla g\|_\infty \lambda \mathrm{e}^\alpha |h|}{w(g)}\Big)^{\frac{1}{p-1}} \Lambda^0(\mathrm{d}w) \bigg)^{p-1}.$$

This implies the second inequality since

$$\bigg(\int_{W_t} \Big(1 + \frac{c}{w(g)}\Big)^{\frac{1}{p-1}} \Lambda^0(\mathrm{d}w) \bigg)^{p-1} \le \bigg\{ 1 + c^{\frac{1}{(p-1)\vee 1}} \gamma_{\rho_0, g}\Big(\frac{1}{p-1}, t\Big)^{(p-1)\wedge 1} \bigg\}^{(p-1)\vee 1}$$

holds for $c \ge 0$ according to Lemma 2.3 and the triangle inequality for the norm

$$\|F\|_{\frac{1}{p-1}} := \bigg(\int_{W_t} |F|^{\frac{1}{p-1}}(w) \Lambda^0(\mathrm{d}w) \bigg)^{(p-1)\wedge 1}, \quad r > 0.$$

□

Next, we consider the log-Harnack inequality.

Theorem 4.2. *Let α, λ and ρ_0, g be in Theorem 4.1. For any positive $f \in \mathscr{B}_b(\mathbb{R}^d)$,*

$$P_t \log f(x+h) \le \log P_t f(x) + \lambda e^{\alpha}|h|\big(\|\nabla \log(\rho_0 g)\|_\infty + \|\nabla g\|_\infty \gamma_{\rho_0,g}(1, t\wedge 1)\big)$$

holds for $t > 0$ and $x, h \in \mathbb{R}^d$.

Proof. Again, due to the semigroup property and the Jensen inequality, it suffices to prove for $t \in (0,1]$. Let

$$\Omega(\mathrm{d}w^1, \mathrm{d}w^0, \mathrm{d}z, \mathrm{d}s) = \frac{g(z)}{w(g)+g(z)}\Lambda^1(\mathrm{d}w^1)\Lambda^0(\mathrm{d}w^0)\nu_0(\mathrm{d}z)\mathrm{d}s,$$

which is a probability measure on $W_t \times W_t \times \mathbb{R}^d \times [0,t]$ according to (3.3) for $g(z,s) = g(z), f = 1$(hence, $h_\varepsilon = 1$) and $\varepsilon = 0$. Let

$$G(w,z,s) = \frac{(w(g)+g(z))(\rho_0 g)(z - \sigma_s^{-1}T_s h)}{(w(g)+g(z-\sigma_s^{-1}T_s h))(\rho_0 g)(z)},$$

which is a probability density w.r.t. Ω by the same reason. Moreover, using $\log f$ to replace f in (3.3) with $\varepsilon = 1$, we have

$$P_t \log f(x+h) = \int_{W_t\times W_t\times\mathbb{R}^d\times[0,t]} (\log h_0)(w^1+w^0+z1_{[s,t]})G(w^0,z,s)\Omega(\mathrm{d}w^1,\mathrm{d}w^0,\mathrm{d}z,\mathrm{d}s).$$

So, by the Young inequality (see [3, Lemma 2.4]) and (3.3) with $\varepsilon = 0$, we obtain

$$\begin{aligned} P_t \log f(x+h) &\le \log \int_{W_t\times W_t\times\mathbb{R}^d\times[0,t]} h_0(w^1+w^0+z1_{[s,t]})\Omega(\mathrm{d}w^1,\mathrm{d}w^0,\mathrm{d}z,\mathrm{d}s) + \Omega(G\log G) \\ &= \log P_t f(x) + \int_{W_t\times\mathbb{R}^d\times[0,t]} \Big\{\frac{(\rho_0 g)(z-\sigma_s^{-1}T_s h)}{w(g)+g(z-\sigma_s^{-1}T_s h)} \log G(w,z,s)\Big\}\Lambda^0(\mathrm{d}w)\mathrm{d}z\mathrm{d}s, \end{aligned}$$

where $\Omega(G\log G)$ is the integral of $G\log G$ w.r.t. the probability measure Ω. Since for $t \in (0,1]$ one has

$$G(w,z,s) \le \exp\left[\left(\frac{\|\nabla g\|_\infty}{w(g)} + \|\nabla \log(\rho_0 g)\|_\infty\right)\lambda e^{\alpha}|h|\right],$$

and since (4.2) and the integral transform $z \mapsto z + \sigma_s^{-1}T_s h$ imply that

$$\begin{aligned} &\int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{(\rho_0 g)(z-\sigma_s^{-1}T_s h)}{w(g)+g(z-\sigma_s^{-1}T_s h)}\Lambda^0(\mathrm{d}w)\mathrm{d}z\mathrm{d}s \\ &= \int_{W_t\times\mathbb{R}^d\times[0,t]} \frac{g(z)}{w(g)+g(z)}\Lambda^0(\mathrm{d}w)\nu_0(\mathrm{d}z)\mathrm{d}s = 1, \end{aligned}$$

we conclude that

$$P_t \log f(x+h) \le \log P_t f(x) + \lambda e^{\alpha}|h|\bigg(\|\nabla \log(\rho_0 g)\|_\infty + \|\nabla g\|_\infty \int_{W_t} \frac{1}{w(g)}\Lambda^0(\mathrm{d}w)\bigg).$$

This completes the proof according to Lemma 2.3. $\square$

Finally, we consider a specific situation for ν having an α-stable like lower bound. Comparing with the Harnack inequality (1.3) derived recently in [39], our result (4.4) is better for small time and small $|h|$, and we only need the specific lower bound in a neighborhood of 0.

Corollary 4.3. *Let A_s and $\|\sigma_s^{-1}\|$ be bounded above, and let $\nu(\mathrm{d}z) \geq h(|z|)\mathrm{d}z$ for some positive decreasing function $h \in C^1((0,\infty))$ such that*

$$\sup_{r>0} \frac{|h'(r)|}{h(r)+h(r)^2} < \infty. \tag{4.3}$$

Then for any $p>1$ there exist two constants $c_1, c_2 > 0$ such that for any positive $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$(P_t f(x+h))^p \leq P_t f^p(x) \mathrm{e}^{c_2|h|}\Big(1 + c_2|h|^{\frac{1}{(p-1)\vee 1}} \int_0^\infty r^{\frac{2-p}{p-1}} \mathrm{e}^{-c_1(t\wedge 1)r(h^{-1}(r))^d}\mathrm{d}r\Big)^{(p-1)\vee 1} \tag{4.4}$$

holds for $t>0, x, h \in \mathbb{R}^d$. Moreover, there exist constants $c_1, c_2 > 0$ such that

$$P_t \log f(x+h) \leq \log P_t f(x) + c_2|h| \int_0^\infty \mathrm{e}^{-c_1(t\wedge 1)r(h^{-1}(r))^d}\mathrm{d}r, \quad x, h \in \mathbb{R}^d, t>0 \tag{4.5}$$

holds for positive $f \in \mathscr{B}_b(\mathbb{R}^d)$.

Proof. Obviously, it suffices to prove for $t \in (0,1]$. Let $\rho_0(z) = h(|z|)$ and $g(z) = \frac{1}{1\vee\rho_0(z)} = \frac{1}{1\vee h(|z|)}$. Then it is easy to see from (4.3) that $\|\nabla \log(\rho_0 g)\|_\infty, \|\nabla g\|_\infty < \infty$. Moreover, since

$$(\rho_0 g)(z) = h(|z|) \wedge 1 = 1, \text{ if } g(z) \leq 1,$$

for $r \geq 1$ we have

$$\nu_0(1-\mathrm{e}^{-rg}) \geq \frac{r}{2}\nu_0(g 1_{\{g \leq r^{-1}\}}) \geq \frac{\kappa(d)r}{2}\int_0^{h^{-1}(r)} s^{d-1}\mathrm{d}s \geq c_1(h^{-1}(r))^d$$

for some constants $c_1 > 0$. Thus, for any $\theta > 1$, there exists constants $c_2 > 0$ such that

$$\gamma_{\rho_0,g}(\theta,t) \leq c_2 \int_0^\infty r^{\theta-1}\exp\big[-tc_1 r(h^{-1}(r))^d\big]\mathrm{d}r$$

holds for $\theta = \frac{p}{p-1}$. Therefore, (4.4) and (4.5) follow from Theorems 4.1 and 4.2 respectively. $\square$

To illustrate the Corollary 4.3, we consider $\nu(\mathrm{d}z) \geq bc_0|z|^{-(d+\alpha)}$ for some $c_0 > 0$ and $\alpha \in (0,2)$. Letting $h(r) = c_0 r^{-(d+\alpha)}$ we have

$$\int_0^\infty r^{\frac{2-p}{p-1}}\mathrm{e}^{-c_1 tr(h^{-1}(r))^d}\mathrm{d}r \leq c' t^{-\frac{d+\alpha}{\alpha(p-1)}}$$

for some constant $c' > 0$ and all $t \in (0, 1]$. Therefore,

$$(P_t f(x+h))^p \le P_t f^p(x) e^{c|h|} (t \wedge 1)^{-\frac{d+\alpha}{\alpha(p-1)}}$$

and

$$P_t \log f(x+h) \le \log P_t f(x) + \frac{c|h|}{(t \wedge 1)^{\frac{\alpha+d}{\alpha}}}$$

hold for all $t > 0, x, h \in \mathbb{R}^d$ and positive $f \in \mathscr{B}_b(\mathbb{R}^d)$. One may also derive explicit Harnack and log-Harnack inequalities for the case that $\nu(\mathrm{d}z) \ge c_0 |z|^{-d} \log^{\varepsilon}(1 + |z|^{-1})$ for some $c_0, \varepsilon > 0$.

4.2 Applications

For applications of our results derived in this section, we introduce some applications of Harnack inequalities which are essentially organized or generalized from [33, 36, 38]. As most results presented below are not yet well known, we include brief proofs for readers' convenience.

Let E be a topology space with Borel σ-field $\mathscr{B}$, let $\mathscr{B}(E)$(resp. $\mathscr{B}_b(E), \mathscr{B}_b^+(E)$) denote the set of all measurable (resp. bounded measurable, bounded non-negative measurable) functions on E, and let $C(E)$ (resp. $C_b(E), C_b^+(E)$) stands for the set of a continuous (resp. bounded continuous, bounded non-negative continuous) functions on E. We recall some notions which will be considered in this subsection.

Definition 4.1. Let μ be a probability measure on $(E, \mathscr{B})$, and let P be a bounded linear operator on $\mathscr{B}_b(E)$.

- (i) μ is called *quasi-invariant* of P, if μP is absolutely continuous w.r.t. μ, where $(\mu P)(A) := \mu(P 1_A),\ A \in \mathscr{F}$. If $\mu P = \mu$ then μ is called an *invariant probability measure* of P.
- (ii) A measurable function p on E^2 is called the *kernel* of P w.r.t. μ, if

$$Pf = \int_E p(\cdot, y) f(y) \mu(\mathrm{d}y), \quad f \in \mathscr{B}_b(E).$$

- (iii) P is called a *Feller* operator, if $PC_b(E) \subset C_b(E)$, while it is called a *strong Feller* operator if $P\mathscr{B}_b(E) \subset C_b(E)$.

From now on, we let P be a Markov operator given by

$$Pf(x) = \int_E f(y) P(x, \mathrm{d}y), \quad f \in \mathscr{B}_b(E), x \in E$$

for a transition probability measure $P(x, \mathrm{d}y)$. We will consider the following general version of *Harnack* type inequality for P:

$$\Phi(Pf(x)) \le \{P\Phi(f)(y)\} e^{\Psi(x,y)}, \quad x, y \in E, f \in \mathscr{B}_b^+(E), \tag{4.6}$$

where Φ is a non-negative function on $[0,\infty)$ and Ψ is a measurable non-negative function on E^2. In particular, the log-Harnack inequality and Harnack inequality with a power $p>1$ addressed above refer to $\Phi(r)=\mathrm{e}^r$ and $\Phi(r)=r^p$ respectively.

Theorem 4.4. *Let μ be a quasi-invariant probability measure of P. Let $\Phi\in C^1([0,\infty))$ be an increasing function with $\Phi'(1)>0$ and $\Phi(\infty):=\lim_{r\to\infty}\Phi(r)=\infty$, such that (4.6) holds.*

(1) *For any $x,y\in E$, $P(x,\cdot)$ and $P(y,\cdot)$ are equivalent.*

(2) *If $\lim_{y\to x}\{\Psi(x,y)+\Psi(y,x)\}=0$ holds for all $x\in E$, then P is strong Feller.*

(3) *P has a kernel p w.r.t. μ, so that any invariant probability measure of P is absolutely continuous w.r.t. μ.*

(4) *P has at most one invariant probability measure and if it has, the kernel of P w.r.t. the invariant probability measure is strictly positive.*

(5) *The kernel p of P w.r.t. μ satisfies*

$$\int_E p(x,\cdot)\Phi^{-1}\Big(\frac{p(x,\cdot)}{p(y,\cdot)}\Big)\mathrm{d}\mu\le\Phi^{-1}(\mathrm{e}^{\Psi(x,y)}),\quad x,y\in E,$$

where $\Phi^{-1}(\infty):=\infty$ by convention.

(6) *If $r\Phi^{-1}(r)$ is convex for $r\ge 0$, then the kernel p of P w.r.t. μ satisfies*

$$\int_E p(x,\cdot)p(y,\cdot)\mathrm{d}\mu\ge\mathrm{e}^{-\Psi(x,y)},\quad x,y\in E.$$

Proof. (1) Let $A\in\mathscr{B}$ be such that $P(y,A)=0$. Applying (4.6) to $n1_A$ we obtain

$$\Phi(nP1_A(x))\le\mathrm{e}^{\Psi(x,y)}P\Phi(n1_A)(y)=\mathrm{e}^{\Psi(x,y)}\Phi(0).$$

Since $\Phi(r)\to\infty$ as $r\to\infty$, letting $n\to\infty$ we conclude that $P(x,A)=P1_A(x)=0$. That is, $P(x,\cdot)$ is absolutely continuous w.r.t. $P(y,\cdot)$ and vice versa.

(2) Let $f\in\mathscr{B}_b(E)$ be positive. Applying (4.6) to $1+\varepsilon f$ in place of f for $\varepsilon>0$, we have

$$\Phi(1+\varepsilon Pf(x))\le\{P\Phi(1+\varepsilon f)(y)\}\mathrm{e}^{\Psi(x,y)},\quad x,y\in E,\varepsilon>0.$$

By the Taylor expansion this implies

$$\Phi(1)+\varepsilon\Phi'(1)Pf(x)+\circ(\varepsilon)\le\{\Phi(1)+\varepsilon\Phi'(1)Pf(y)+\circ(\varepsilon)\}\mathrm{e}^{\Psi(x,y)}\tag{4.7}$$

for small $\varepsilon>0$. Letting $y\to x$ we obtain

$$\varepsilon Pf(x)\le\varepsilon\liminf_{y\to x}Pf(y)+\circ(\varepsilon).$$

Thus, $Pf(x) \le \lim_{y\to x} Pf(y)$ holds for all $x \in E$. On the other hand, letting $x \to y$ in (4.7) gives $Pf(y) \ge \limsup_{x\to y} Pf(x)$ for any $y \in E$. Therefore, Pf is continuous.

(3) To prove the existence of kernel, it suffices to prove that for any $A \in \mathscr{F}$ with $\mu(A) = 0$ we have $P1_A \equiv 0$. Applying (4.6) to $f = 1 + n1_A$, we obtain

$$\Phi(1 + nP1_A(x)) \int_E \mathrm{e}^{-\Psi(x,y)} \mu(\mathrm{d}y) \le \int_E \Phi(1 + n1_A)(y)(\mu P)(\mathrm{d}y), \quad n \ge 1. \tag{4.8}$$

Since $\mu(A) = 0$ and μ is quasi-invariant for P, we have $1_A = 0, \mu P$-a.s. So, it follows from (4.8) that

$$\Phi(1 + nP1_A(x)) \le \frac{\Phi(1)}{\int_E \mathrm{e}^{-\Psi(x,y)} \mu(\mathrm{d}y)} < \infty, \quad x \in E, n \ge 1.$$

Since $\Phi(1 + n) \to \infty$ as $n \to \infty$, this implies that $P1_A(x) = 0$ for all $x \in E$.

Now, for any invariant probability measure μ_0 of P, if $\mu(A) = 0$ then $P1_A \equiv 0$ implies that $\mu_0(A) = \mu_0(P1_A) = 0$. Therefore, μ_0 is absolutely continuous w.r.t. μ.

(4) We first prove that the kernel of P w.r.t. an invariant probability measure μ_0 is strictly positive. To this end, it suffices to show that for any $x \in E$ and $A \in \mathscr{F}$, $P1_A(x) = 0$ implies that $\mu_0(A) = 0$. Since $P1_A(x) = 0$, applying (4.6) to $f = 1 + nP1_A$ we obtain

$$\Phi(1 + nP1_A(y)) \le \{P\Phi(1 + n1_A)(x)\}\mathrm{e}^{\Psi(y,x)} = \Phi(1)\mathrm{e}^{\Psi(y,x)}, \quad y \in E, n \ge 1.$$

Letting $n \to \infty$ we conclude that $P1_A \equiv 0$ and hence, $\mu_0(A) = \mu_0(P1_A) = 0$.

Next, let μ_1 be another invariant probability measure of P, by (2) we have $\mathrm{d}\mu_1 = f\mathrm{d}\mu_0$ for some probability density function f. We aim to prove that $f = 1, \mu_0$-a.e. Let $p(x, y) > 0$ be the kernel of P w.r.t. μ_0, and let $P^*(x, \mathrm{d}y) = p(y, x)\mu_0(\mathrm{d}y)$. Then

$$P^*g = \int_E g(y)P^*(\cdot, \mathrm{d}y), \quad g \in \mathscr{B}_b(E)$$

is the adjoint operator of P w.r.t. μ_0. Since μ_0 is P-invariant, we have

$$\int_E gP^*1 \,\mathrm{d}\mu_0 = \int_E Pg \,\mathrm{d}\mu_0 = \int_E g \,\mathrm{d}\mu_0, \quad g \in \mathscr{B}_b(E).$$

This implies that $P^*1 = 1, \mu_0$-a.e. Thus, for μ_0-a.e. $x \in E$ the measure $P^*(x, \cdot)$ is a probability measure. On the other hand, since μ_1 is P-invariant, we have

$$\int_E (P^*f)g \,\mathrm{d}\mu_0 = \int_E fPg \,\mathrm{d}\mu_0 = \int_E Pg \,\mathrm{d}\mu_1 = \int_E g \,\mathrm{d}\mu_1 = \int_E fg \,\mathrm{d}\mu_0, \quad g \in \mathscr{B}_b(E).$$

This implies that $P^*f = f, \mu$-a.e. Therefore, for any $r > 0$ we have

$$\int_E P^* \frac{1}{f+1} \,\mathrm{d}\mu_0 = \int_E \frac{1}{f+1} \,\mathrm{d}\mu_0 = \int_E \frac{1}{P^*f+1} \,\mathrm{d}\mu_0.$$

When $P^*(x,\cdot)$ is a probability measure, by the Jensen inequality one has $P^*\frac{1}{1+f}(x) \geq \frac{1}{P^*f+1}(x)$ and the equation holds if and only if f is constant $P^*(x,\cdot)$-a.s. Hence, f is constant $P^*(x,\cdot)$-a.s. for μ_0-a.e. x. Since $p(x,y)>0$ for any $y \in E$ such that μ_0 is absolutely continuous w.r.t. $P^*(x,\cdot)$ for any $x \in E$, we conclude that f is constant μ_0-a.s. Therefore, $f = 1$ μ_0-a.s. since f is a probability density function.

(5) Applying (4.6) to

$$f = n \wedge \Phi^{-1}\Big(\frac{p(x,\cdot)}{p(y,\cdot)}\Big)$$

and letting $n \to \infty$, we obtain the desired inequality.

(6) Let $r\Phi^{-1}(r)$ be convex for $r \geq 0$. By the Jensen inequality we have

$$\int_E p(x,\cdot)\Phi^{-1}(p(x,\cdot))\mathrm{d}\mu \geq \Phi^{-1}(1).$$

So, applying (4.6) to

$$f = n \wedge \Phi^{-1}(p(x,\cdot))$$

and letting $n \to \infty$, we obtain

$$\int_E p(x,\cdot)p(y,\cdot)\mathrm{d}\mu \geq \mathrm{e}^{-\Psi(x,y)}\Phi\bigg(\int_E p(x,\cdot)\Phi^{-1}(p(x,\cdot))\mathrm{d}\mu\bigg) \geq \mathrm{e}^{-\Psi(x,y)}.$$

□

Next, by Theorem 4.4(1), if (4.6) holds then

$$p_{x,y}(z) := \frac{P(x,\mathrm{d}z)}{P(y,\mathrm{d}z)}$$

exists. We aim to describe this function using the Harnack inequality. For simplicity, we only consider the Harnack inequality with a power $p > 1$

$$(Pf(x))^p \leq (Pf^p(y))\mathrm{e}^{\Psi(x,y)}, \quad x,y \in E, f \in \mathscr{B}_b^+(E) \tag{4.9}$$

and the log-Harnack inequality

$$P(\log f)(x) \leq \log Pf(y) + \Psi(x,y), \quad x,y \in E, f \geq 1, f \in \mathscr{B}_b(E). \tag{4.10}$$

Proposition 4.5. *(4.9) holds if and only if $p_{x,y}$ exists and satisfies*

$$P\big\{p_{x,y}^{1/(p-1)}\big\}(x) \leq \Psi(x,y)^{1/(p-1)}, \quad x,y \in E; \tag{4.11}$$

while (4.10) holds if and only if $p_{x,y}$ exists and satisfies

$$P\{\log p_{x,y}\}(x) \leq \Psi(x,y), \quad x,y \in E. \tag{4.12}$$

Proof. (1) Applying (4.9) to $f_n(z) := \{n \wedge p_{x,y}(z)\}^{1/(p-1)},\ n \geq 1$, we obtain

$$\begin{aligned}(Pf_n(x))^p &\leq \Psi(x,y)Pf_n^p(y) = \Psi(x,y)\int_E \{n \wedge p_{x,y}(z)\}^{p/(p-1)}P(y,\mathrm{d}z)\\ &\leq \Psi(x,y)\int_E \{n \wedge p_{x,y}(z)\}^{1/(p-1)}P(x,\mathrm{d}z) = \Psi(x,y)Pf_n(x).\end{aligned}$$

Thus,

$$P\{p_{x,y}^{1/(p-1)}\}(x) = \lim_{n\to\infty} Pf_n(x) \leq \Psi(x,y)^{1/(p-1)}.$$

So, (4.9) implies (4.11).

On the other hand, if (4.11) holds then for any $f \in \mathscr{B}_b^+(E)$, by the Hölder inequality

$$\begin{aligned}Pf(x) &= \int_E \{p_{x,y}\}(z)f(z)P(y,\mathrm{d}z) \leq (Pf^p(y))^{1/p}\bigg(\int_E p_{x,y}(z)^{p/(p-1)}P(y,\mathrm{d}z)\bigg)^{(p-1)/p}\\ &= (Pf^p(y))^{1/p}(Pp_{x,y}^{1/(p-1)}(x))^{(p-1)/p} \leq (Pf^p(y))^{1/p}\Psi(x,y)^{1/p}.\end{aligned}$$

Therefore, (4.9) holds.

(2) We shall use the following *Young inequality*: for any probability measure ν on E, if $g_1, g_2 \geq 0$ with $\nu(g_1) = 1$, then

$$\nu(g_1g_2) \leq \nu(g_1\log g_1) + \log\nu(\mathrm{e}^{g_2}).$$

For $f \geq 1$, applying the above inequality for $g_1 = p_{x,y}, g_2 = \log f$ and $\nu = P(y,\cdot)$, we obtain

$$\begin{aligned}P(\log f)(x) &= \int_E \{p_{x,y}(z)\log f(z)\}P(y,\mathrm{d}z)\\ &\leq P(\log p_{x,y})(x) + \log Pf(y).\end{aligned}$$

So, (4.12) implies (4.10). On the other hand, applying (4.10) to $f_n = 1 + np_{x,y}$, we arrive at

$$\begin{aligned}&P\{\log p_{x,y}\}(x) \leq P(\log f_n)(x) - \log n\\ &\leq \log Pf_n(y) - \log n + \Psi(x,y) = \log\frac{n+1}{n} + \Psi(x,y).\end{aligned}$$

Therefore, by letting $n \to \infty$ we obtain (4.12). $\square$

Finally, we consider the hyperbounded property and the entropy-cost inequality implied by (4.9) and (4.10). Let P have an invariant probability measure μ. Then $\|\cdot\|_{p\to q}$ stands for the operator norm from $L^p(\mu)$ to $L^q(\mu)$. Moreover, for a non-negative measurable function Ψ on $E\times E$, and for $\mathscr{C}(\nu,\mu)$ the class of all couplings of μ and ν, let

$$W_\Psi(\mu,\nu) = \inf_{\pi\in\mathscr{C}(\nu,\mu)}\int_{E\times E}\Psi(x,y)\pi(\mathrm{d}x,\mathrm{d}y)$$

be the transportation-cost from ν to μ induced by the cost-function Ψ.

Proposition 4.6. *Let P have an invariant probability measure μ.*

(1) *(4.9) implies*

$$\|P\|_{p\to\delta p} \le \int_E \frac{\mu(\mathrm{d}x)}{\left\{\int_E \exp[-\Psi(x,y)]\mu(\mathrm{d}y)\right\}^{\delta}}, \quad \delta > 1.$$

(2) *Let P^* be the adjoint operator of P in $L^2(\mu)$. Then (4.10) implies*

$$\int_E (P^* f)\log P^* f \mathrm{d}\mu \le W_\Psi(f\mu,\mu), \quad f\ge 0, \int_E f\mathrm{d}\mu = 1.$$

Proof. (1) Let $f\in \mathscr{B}_b^+(E)$ be such that $\int_E f^p\mathrm{d}\mu = 1$. By (4.9) we have

$$(Pf)^p \mathrm{e}^{-\Psi(x,y)} \le Pf^p(y).$$

Then integrating both sides w.r.t. $\mu(\mathrm{d}y)$ we obtain

$$(Pf)^p(x)\int_E \mathrm{e}^{-\Psi(x,y)}\mu(\mathrm{d}y) \le 1.$$

This implies that

$$\int_E (Pf)^{\delta p}\mathrm{d}\mu \le \int_E \frac{\mu(\mathrm{d}x)}{\left\{\int_E \exp[-\Psi(x,y)]\mu(\mathrm{d}y)\right\}^{\delta}}, \quad \delta > 1.$$

(2) Let $f\ge 0$ be such that $\int_E f\mathrm{d}\mu = 1$. Applying (4.10) to P^*f in place of f, we obtain

$$P(\log P^* f)(x) \le \log(PP^*f)(y) + \Psi(x,y).$$

Integrating both sides w.r.t. $\pi\in\mathscr{C}(f\mu,\mu)$, using the Jensen inequality and noting that μ is PP^*-invariant, we arrive at

$$\begin{aligned}
&\int_E (P^*f)\log P^* f\mathrm{d}\mu = \int_{E\times E} (P\log P^* f)(x)\pi(\mathrm{d}x,\mathrm{d}y)\\
&\le \int_{E\times E} \log(PP^*f)(y)\pi(\mathrm{d}x,\mathrm{d}y) + \int_{E\times E}\Psi\mathrm{d}\pi\\
&\le \log\int_{E\times E}(PP^*f)(y)\pi(\mathrm{d}x,\mathrm{d}y) + \int_{E\times E}\Psi\mathrm{d}\pi = \int_{E\times E}\Psi\mathrm{d}\pi.
\end{aligned}$$

This completes the proof by minimizing in $\pi\in\mathscr{C}(f\mu,\mu)$. □

Acknowledgement. The author would like to thank the referee and Dr. Jian Wang for corrections and useful comments.

References

[1] D. Applebaum, *Lévy Processes and Stochastic Calculus,* Cambridge University Press, 2004.

[2] M. Arnaudon, A. Thalmaier, F.-Y. Wang, *Harnack inequality and heat kernel estimates on manifolds with curvature unbounded below,* Bull. Sci. Math. 130(2006), 223–233.

[3] M. Arnaudon, A. Thalmaier, F.-Y. Wang, *Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds,* Stoch. Proc. Appl. 119(2009), 3653–3670.

[4] D. Bakry, *On Sobolev and logarithmic Sobolev inequalities for Markov semigroups,* "New Trends in Stochastic Analysis" (Editors: K. D. Elworthy, S. Kusuoka, I. Shigekawa), Singapore: World Scientific, 1997.

[5] J. M. Bismut, *Large Deviations and the Malliavin Calculus,* Boston: Birkhäuser, MA, 1984.

[6] V. I. Bogachev, M. Röckner, B. Schmuland, *Generalized Mehler semigroups and applications,* Probab. Th. Relat. Fields 105(1996), 193–225.

[7] K. Bogdan, T. Jakubowski, *Estimate of heat kernel of fractional Laplacian perturbed by gradient operators,* Comm. Math. Phys. 271(2007), 179–198.

[8] B. Böttcher, R. L. Schilling, J. Wang, *Constructions of coupling processes for Lévy processes,* Stoch. Proc. Appl. 121(2011), 1201–1216.

[9] G. Da Prato, M. Röckner, F.-Y. Wang, *Singular stochastic equations on Hilbert spaces: Harnack inequalities for their transition semigroups,* J. Funct. Anal. 257 (2009), 992–017.

[10] K.D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups,* J. Funct. Anal. 125(1994), 252–286.

[11] M. Gordina, M. Röckner, F.-Y. Wang, *Harnack inequalities for subordinate semigroups,* to appear in Potential Analysis. arXiv: 1004.3016.

[12] A. Guillin, F.-Y. Wang, *Degenerate Fokker-Planck equations : Bismut formula, gradient estimate and Harnack inequality,* preprint.

[13] N. Jacob, *Pseudo Differential Operators and Markov Processes (Volume I),* Imperial College Press, London, 2001.

[14] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, (1987), Berlin.

[15] H. Kawabi, *The parabolic Harnack inequality for the time dependent Ginzburg-Landau type SPDE and its application,* Pot. Anal. 22(2005), 61–84.

[16] V. Knopova, R.L. Schilling, *A note on the existence of transition probability densities for Lévy processes,* to appear in Forum Math. arXiv:1003.1419.

[17] V. Knopova, R.L. Schilling, *Transition density estimates for a class of Lévy and Lévy-type processes,* to appear in J. Theory Probab. arXiv:0912.1482v1.

[18] W. Liu, F.-Y. Wang, *Harnack inequality and strong Feller property for stochastic fast diffusion equations,* J. Math. Anal. Appl. 342(2008), 651–662.

[19] J. Mecke, *Stationaire Zufällige Maße auf lokalkompakten abelschen Gruppen,* Z. Wahrsch. verw. Geb. 9(1967), 36–58.

[20] S.-X. Ouyang, *Harnack inequalities and applications for multivalued stochastic evolution equations,* to appear in Infin. Dimens. Anal. Quant. Probab. Relat. Topics, accessible on arXiv:0908.3630

[21] S.-X. Ouyang, M. Röckner, F.-Y. Wang, *Harnack inequalities and applications for Ornstein-Uhlenbeck semigroups with jump,* to appear in Pot. Anal. arXiv: 0908.2889.

[22] E. Priola, J. Zabczyk, *Densities for Ornstein-Uhlenbeck processes with jumps,* Bull. Lond. Math. Soc. 41(2009), 41–50.

[23] M. Röckner, *Stochastic analysis on configuration spaces: basic ideas and recent results,* New Directions in Dirichlet Forms, 157–231. AMS/IP Stud. Math. 8, Amer. Math. Soc., Providence, RI, 1998.

[24] M. Röckner, F.-Y. Wang, *Harnack and functional inequalities for generalized Mehler semigroups,* J. Funct. Anal. 203(2003), 237–261.

[25] M. Röckner, F.-Y. Wang, *Log-Harnack inequality for stochastic differential equations in Hilbert spaces and its consequences,* Infin. Dimens. Anal. Quant. Probab. Relat. Topics 13(2010), 27–37.

[26] R. L. Schilling, J. Wang, *On the coupling property of Lévy processes,* Inst. Henri Poinc. Probab. Stat. arXiv:1006.5288

[27] R. L. Schilling, P. Sztonyk, J. Wang, *Coupling property and gradient estimates of Lévy processes via the symbol,* to appear in Bernoulli. arXiv:1011.1067

[28] R. L. Schilling, *Subordination in the sense of Bochner and a related functional caculus,* J. Austral. Math. Soc. Ser. A 64(1998), 368–396.

[29] R. Song, Z. Vondraček, *Potential theory of subordinate Brownian motion,* in "Potential Analysis of Stable Processes and its Extensions", 87-176, Lecture Notes in Math. Vol. 1980, Springer, Berlin, 2009.

[30] A. Takeuchi, *Bismut-Elworthy-Li-Type formula for stochastic differential equations with jumps,* J. Theory Probab. 23(2010), 576–604.

[31] F.-Y. Wang, *On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups,* Probab. Theory Relat. Fields 108(1997), 87–101.

[32] F.-Y. Wang, *Harnack inequality and applications for stochastic generalized porous media equations,* Ann. Probab. 35(2007), 1333–1350.

[33] F.-Y. Wang, *Harnack inequalities on manifolds with boundary and applications,* J. Math. Pures Appl. 94(2010), 304–321.

[34] F.-Y. Wang, *Gradient estimate for Ornstein-Uhlenbeck jump processes,* Stoch. Proc. Appl. 121(2011), 466–478.

[35] F.-Y. Wang, *Coupling for Ornstein-Uhlenbeck Processes with Jumps,* to appear in Bernoulli. arXiv: 1002.2890.

[36] F.-Y. Wang, L. Xu, *Derivative formula and applications for hyperdissipative stochastic Navier-Stokes/Burgers equations,* to appear in Infin. Dimens. Anal. Quant. Probab. Relat. Topics, accessible on arXiv:1009.1464.

[37] F.-Y. Wang, J.-L. Wu, L. Xu, *Log-Harnack inequality for stochastic Burgers equations and applications,* to appear in J. Math. Anal. Appl. arXiv:1009.5948

[38] F.-Y. Wang, C. Yuan, *Harnack inequalities for functional SDEs with multiplicative noise and applications,* Stoch. Proc. Appl. 121(2011), 2692–2710.

[39] J. Wang, *Harnack inequalities for Ornstein-Uhlenbeck processes driven by Lévy processes,* preprint.

[40] T.-S. Zhang, *White noise driven SPDEs with reflection: strong Feller properties and Harnack inequalities,* Potential Anal. 33 (2010),137–151.